Foundation of Cryptography, Spring 2014 Iftach Haitner

Problem set 2

March 26, 2014 Due— In class: April 8. Via Email: April 10

- Please submit the handout in class, or email me, in case you write in LaTeX
- Write clearly and shortly using sub claims if needed. The emphasize in most questions is on the proofs (no much point is writing a “solution" w/o proving its correctness)
- For Latex users, a solution example can be found in the course web site.
- In it ok to work in (small) groups, but please write the id list of your partners in the solution file, and each student should write his solution by *himself* (joint effort is only allowed in the “thinking phase")
- The notation we use appear in the introduction part of the first lecture (*Notation* section).

1. A function family is a *non-adaptive* PRF, if it is a PRF according to the definition given in class, but its security should only holds against *non-adaptive* distinguishers: distinguishers that choose all queries to the oracle *before* making the first query (alternatively, they make all there queries at once).

   Assume OWF exists, prove there exists a non-adaptive PRF that is *not* an (adaptive) PRF.

2. (a) Let X_1 and X_2 be jointly distributed random variables over domain $\mathcal{U}$, and let D_1 and D_2 be their marginal distributions, respectively.
   Prove that $\mathrm{SD}(D_1, D_2) \leq \Pr[X_1 \neq X_2]$.

   (b) For $k \in \mathbb{N}$, let P_k be the distribution (over $\{0,1\}^k$) induced by concatenating k unbiased independent coins, i.e., each coin is taking the value 1 with probability $\frac{1}{2}$ and 0 otherwise. Let Q_k be the distribution induced by concatenating k ε-biassed independent coins, i.e., each coin is taking the value 1 with probability $\frac{1}{2} + \varepsilon$ and 0 otherwise. We would like to bound the statistical distance between P_k and Q_k.

      i. Describe a process S (i.e., an algorithm) that outputs a pair of values (x, y), and let (X, Y) be the random variable describing the output of S in a random execution. Prove that the marginal distribution of X is P_k, and that the marginal distribution of Y is Q_k.
      Hint: consider the following method for sampling an ε-biassed coin: with probability 2ε output 1, and otherwise, output an unbiased coin.

      ii. Bound the probability that $X \neq Y$.

      iii. Use the first part of the question to bound $\mathrm{SD}(P_k, Q_k)$.

3. Let $\mathcal{F} = \{\mathcal{F}_n = \{f \colon \{0,1\}^n \mapsto \{0,1\}^n\}\}_{n\in\mathbb{N}}$ be a PRF, and let $\mathcal{H} = \{\mathcal{H}_n = \{h \colon \{0,1\}^{2n} \mapsto \{0,1\}^n\}\}_{n\in\mathbb{N}}$ be an efficient pairwise-independent function family.[1] We would like to prove that the function family ensemble $\mathcal{F} \circ \mathcal{H} = \{\mathcal{F}_n \circ \mathcal{H}_n = \{f \circ h \colon f \in \mathcal{F}_n, h \in \mathcal{H}_n\}\}_{n\in\mathbb{N}}$ is a PRF mapping strings of length $2n$ to string of length n.[2]

   (a) Prove that function family ensemble $\{G_n = \Pi_n \circ \mathcal{H}_n\}_{n\in\mathbb{N}}$ is computationally indistinguishable (actually also statistically) indistinguishable from $\{\Pi_{2n,n}\}_{n\in\mathbb{N}}$.
   Do the above using the following proof methodology. Fix an q-query, deterministic oracle-aided algorithm A, and $n \in \mathbb{N}$.

      i. Describe a process S (i.e., an algorithm) that outputs a pair of values (x, y), and let (X, Y) be the random variable describing the output of S in a random execution. Show that X describes the view of A^g, for $g \leftarrow G_n$ (i.e., X lists the query/anmswer pairs in a an execution of A^g for a random $g \leftarrow G_n$). Similarly, show that Y describes the view of $\mathsf{A}^{\pi_{2n}}$, for $\pi_{2n} \leftarrow \Pi_{2n,n}$.

      ii. Bound the probability that $X \neq Y$. This is the hardest part of the question, and only doable if you have defined S above in a suitable way...

[1] Namely, the family $\mathcal{H}_n$, for each $n \in \mathbb{N}$, is pairwise independent.

[2] The symbol $\circ$ stands for function composition, e.g., $f \circ h(x) = f(h(x))$.

Start with proving the claim assuming that A never makes a colliding query: $\mathsf{A}^{g=\pi\circ h}$ never makes two distinct queries x, x' with $h(x) = h(x')$.

iii. Use question (2) to bound the advantage that A has is distinguishing G_n from $\Pi_{2n,n}$.

iv. Prove that for $q \in \mathrm{poly}$, no q-query algorithm (even a randomized one) distinguishes $\{G_n\}_{n\in\mathbb{N}}$ from $\{\Pi_{2n,n}\}_{n\in\mathbb{N}}$ with more than negligible advantage.

(b) Use the above to prove that $\mathcal{F} \circ \mathcal{H}$ is a PRF.

4. The question is about the signature scheme described in Construction 32, of lecture 5 (page 34).

   Does the construction remain secure when $\boldsymbol{r}$ is set to $g(m)$ (rather than to $\pi(h(m))_{1,\ldots,n}$), where g is part of the singing key, and chosen at random by Gen' from a family of pair-wise independent hash functions from $\{0,1\}^*$ to $\{0,1\}^n$?